



Aalborg Universitet

AALBORG UNIVERSITY
DENMARK

Evaluating the Effects of Co-Production Initiatives in Public Service Organizations

Brix, Jacob; Krogstrup, Hanne Kathrine; Mortensen, Nanna Møller

Published in:
Proceedings of the XVIII Nordic Political Science Congress

Publication date:
2017

Document Version
Accepted author manuscript, peer reviewed version

[Link to publication from Aalborg University](#)

Citation for published version (APA):
Brix, J., Krogstrup, H. K., & Mortensen, N. M. (2017). Evaluating the Effects of Co-Production Initiatives in Public Service Organizations. *Proceedings of the XVIII Nordic Political Science Congress*.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy

If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

Evaluating the Effects of Co-Production Initiatives in Public Service Organizations

Jacob Brix, Hanne Kathrine Krogstrup
and Nanna Møller Mortensen*

Aalborg University,
Department of Learning and Philosophy,
Research Group for Evaluation & Capacity Building

**Corresponding author: nannamm@learning.aau.dk*

KEYWORDS

Evaluation; Collaborative Evaluation; Co-production; Theory of Change; New Public Governance; Innovation

ABSTRACT

A change from New Public Management to New Public Governance (NPG) does not occur overnight. This forces public service organizations to develop new hybrid organizational forms as strategic response to the current situation. In NPG the basic assumption is that co-production will result in increased efficiency and effectiveness for public service organizations as a new organizational recipe. However, a recent review determines that only few empirical studies document these claimed effects. To enable the creation of more empirical evidence that establish the effects of co-production, the purpose of our study is to discuss how the outcomes of co-production initiatives can be evaluated. The goal of our study is to inspire scholars and practitioners to evaluate the outcomes of co-production and not least to document these, so a critical mass of data can be created that can be used to substantiate *if* and *how* the new organizational recipe 'co-production' is delivering what we are told – and what we hope for.

INTRODUCTION

The public sector is currently facing a shift from New Public Management to New Public Governance. According to Gouillart and Hallett (2015) the hitherto used management methods have served their purpose of bringing the cost structure in public service organizations more in line the private sector logic. Moreover, there is growing evidence that public sector organizations have harvested most of the proverbial low-hanging fruits (see also Dixon and Hood, 2013). There is hence a need for major organizational innovation (OECD, 2005) in the entire public sector – a transformation is required that has large impact on how present and future public service organizations are to be re-designed and managed (Levy, 2010; MandagMorgen, 2013; Jensen and

Krogstrup, 2017). As response to the requirement for this transformation, collaboration with citizens is increasingly emerging and spreading across the public sector as an important approach to innovate public service delivery. This approach is labelled co-production (e.g. Voorberg et al. 2013), and it seems that co-production itself has turned into a new organizational recipe that serves as response to the transformation challenge (see also OECD, 2011). According to OECD (2011) co-production is an umbrella-phenomenon covering sub-processes such as co-creation; co-design, co-decide and co-evaluation. The results of co-production are claimed to be beneficial both to the individual citizens as part of their empowerment and Capacity Building, and on an organizational level by increasing efficiency and effectiveness of the service provided (Osborne et al. 2016; OECD 2011; Krogstrup and Mortensen, 2017). However, the introduction of co-production as the new organizational recipe has a fundamental weakness that needs to be addressed: only few empirical studies demonstrate the immediate and long-term effects of co-production initiatives (Voorberg, Bekkers, & Tummers, 2013). Löffler (2009) e.g. establishes that “*Evaluations of co-production approaches are rare*” (Löffler, 2009: 11), which we also find strong indications for in our manuscript. The existing empirical studies of co-production show that often the objectives for implementing co-production are not explicitly formalized, and therefore it seems like co-production is considered a value itself for the organizations adopting the new recipe. Due to the lack of objectives it is difficult to link the specific outcomes to the specific goals and hence to evaluate the effects of the new recipe. Voorberg et al. conclude “ (...) *that hardly any empirical material can be found that systematically assess the outcomes of co-production (...)*” (Voorberg et al., 2013, 35). OECD argues along with this, that there is insufficient evidence for how co-production works and therefore how it impacts the effectiveness of the service provided (OECD 2011: 19). They call for further assessment on how to re-balancing co-production in the delivery of public services in order to “ (...) *quantify the potential savings and assess any unintended consequences (...)*” (OECD 2011: 12).

The link between the claimed outcomes of co-production and the empirical studies documenting the actual outcomes is henceforth weak. The purpose of this study is consequently to discuss and propose how prospective co-production initiatives can be evaluated, so empirical evidence in the future can strengthen the link. The goal of our study is hence to deduce a logic for how to evaluate co-production initiatives in public sector organizations. The study has two key contributions. First, by taking point of departure in classical innovation theory, we define three exemplar empirical

arenas in which co-production can occur: in the arena of ‘exploration’, ‘exploitation’ and in the ‘ambidextrous’ arena (March, 1991; Raisch and Birkenshaw, 2008). We discuss and link evaluation theory to the most complex arena – the ambidextrous. This leads to the second key contribution: we discuss that contribution analysis might be the answer for how to evaluate the outcomes of co-production initiatives. However, since empirical evidence are scarce concerning co-production, we cannot approach the contribution analysis in a traditional manner, since it is based on the evidence, we lack. Therefore, we suggest that collaborative evaluation according to Shulha et al. (2016) can act as a scaffolding approach to generate the minimum viable data for the evaluation to be followed through. In line with this, we propose that the eight principles of collaborative evaluation can be utilized as a method to respond to OECD’s (2011) request for increased knowledge about co-evaluation.

The study proceeds as follows. First we discuss and define what co-production is and we exemplify different empirical arenas in which co-production occurs. Then we discuss evaluation theory and link this to the co-production literature to identify similar criteria that enable scholars and practitioners to select the appropriate evaluation method to create the empirical evidence we urgently need. Finally, we provide the implications and the conclusion of the study.

LITERATURE REVIEW

Defining co-production

The concept of co-production was originally developed by Elinor Ostrom and her research group in the 1970s (Ostrom et al. 1978). They used the term co-production to describe the empirical phenomena of citizens contributing to public service delivery. In their study they found positive effects of the service outcome when citizens became ‘co-producers’ (Ostrom et al. 1978). Since the foundational work of co-production was made, later developments emerged. In the 1990s, Ostrom (1996) explained co-production concept as: “(...) *the potential relationships that could exist between the ‘regular’ producer (street-level police officers, school teachers, or health workers) and ‘clients’ who want to be transformed into safer, better educated, educated, or healthier persons.*” (1996, 1079). Inspired by Ostrom’s work the co-production literature has since then taken point of departure in the *potential relationship* between regular producer and the service-using citizens. E.g. Park et al. (1981) builds on Ostrom’s starting point of the potential relationship and define co-

production as: “(...) *the mix of activities that both public service agents and citizens contribute to the provision of public services. The former are involved as professionals, or ‘regular producers’, while ‘citizen production’ is based on voluntary efforts by individuals and groups to enhance the quality and/or quantity of the services they use*” (Parks et al. 1981). This definition has also been adopted in the studies by e.g. Brudney and England 1983, Verschuere, Brandsen and Pestoff 2012, Pestoff 2015, Jacobsen and Andersen 2013, Tuurnas 2016).

To operationalize this potential relationship and the study of co-production we rely on OECD’s (2011) view on co-production as a (group of) units of analysis such as co-design, co-review and not least co-evaluation. Inspired by OECD (2011) and Bovaird and Loeffler (2012), we define co-production as *a partnership between public service employees and citizens, where the citizen is directly involved in designing, planning, delivering and evaluating public services*. This definition includes both an individual and a group perspective to co-production (Brudney and England, 1983). In empirical contexts the potential relationship between public service employees and citizens cannot be isolated from its external and internal context. External contexts are e.g. legislation, technological advance and people in the citizens’ personal network; whereas internal contexts are e.g. leadership and management at the local service organization (Stringer 2015; Krogstrup and Mortensen: 2017). The chief argument for defining co-production as such is elaborated in the section below, where we present how and where co-production has been studied.

Recent studies of co-production

Co-production has been advanced by multiple scholars – especially the discussion of differences and similarities to co-creation are widely published (see e.g. Voorberg et al. 2013; Torfing, Sørensen and Røiseland, 2016). Since our definition presented above clarifies our research locus, we do not delve further into this discussion. Instead we proceed to review recent publications treading co-production. A recent review by Voorberg et al. (2014) determines that fifty per cent of the empirical studies they reviewed in the context of co-production treat the phenomena *co-implementing*¹. Hence, these studies do according to Bason’s (2010) perspective treat co-production as being dependent on shared responsibility for service delivery between the citizen and the (local) government – but where the citizens’ capacity to take active part in the process influences the

¹ Co-implementing is understood by Voorberg et al. (2014) as a collaboration in which the citizen only performs pre-defined tasks by the public service provider.

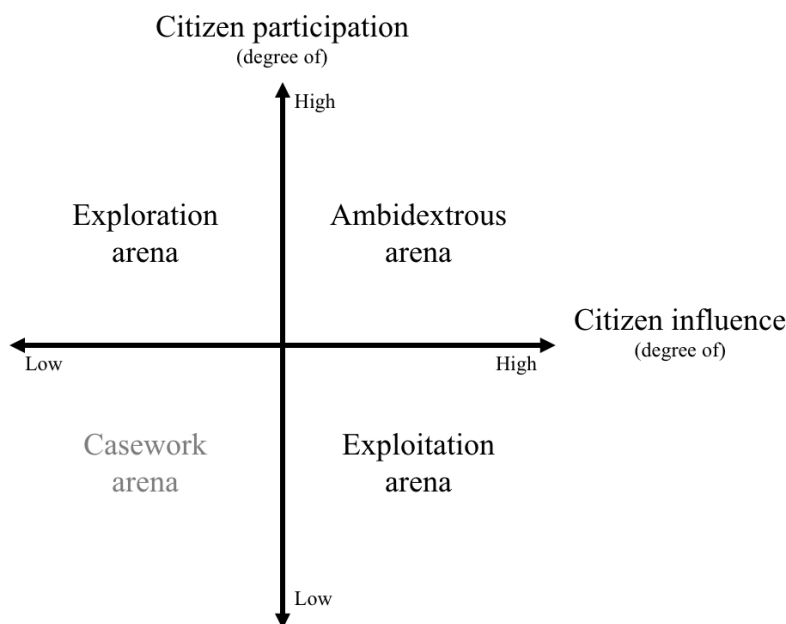
outcome and the experienced outcome of the total service delivery (see also Jacobsen and Andersen, 2013; Pestoff, 2015; Krogstrup and Mortensen, 2017). This view on co-production represents only one aspect of the process according to our definition. We argue, that it is important to broaden the understanding of the citizens' participation in the co-production process, since the co-production process and its sub-processes (co-design, co-planning, co-delivery, and co-evaluation) represent an actual *partnership* with the public service provider. This argument is based on Needham and Carr (2009) and Boyle and Harris (2009) since, they emphasize the importance of power-sharing in the person-centered relationship (Needham & Carr, 2009). In the literature many matrix-figures are developed to explain the contexts in which co-production is studied. Here scholars have conceptualized co-production contexts by dividing the process into researchable typologies by differentiating between varied types and degrees of interactions between citizens and the public service providers (e.g. Bovarid and Loeffler, 2012; Pestoff, 2015). Krogstrup and Mortensen (2017) expand this view by adopting a Capacity Building approach by conceptualizing a framework to illustrate the organizational perspective and the degree of the citizens' involvement. Since these different views exist on the partnership between public service organizations and citizens that co-produce the service they are referred to, it is a daunting task to evaluate processes and the effects of co-production based on the theoretical insights about co-production. Even though it is a daunting task to evaluate the effects of co-production (OECD, 2011) we cannot allow this to be an excuse for not trying to do so. In stead of basing evaluation methodologies on theoretical abstractions of co-production, we may need to base our evaluation models on the empirical contexts in which co-production initiatives occur. This we claim will enable us to determine relevant tools that fit the purpose and goal of co-evaluating the outcomes of co-production initiatives, so empirical evidence can be made in the future to establish the functioning and non-functioning mechanisms that lead to the desired effects (OECD, 2011; Voorberg et al. 2014).

Relating co-production to different public service contexts

To understand how to contextualize the co-production methodologies to different public service contexts we are inspired by classical innovation theory. Here the tension between exploration and exploitation, as well as ability to maneuver both, organizational ambidexterity, have been studied for decades (see e.g. Raisch and Birkinshaw, 2008; O'Reilly and Tushman, 2013). The original dichotomy between exploration and exploitation made popular by James March states that *exploitation* is about efficiency, control, certainty and reduction of variety, whereas *exploration* is

about innovation as proxied by search, discovery, and autonomy (see e.g. March, 1991). *Organizational Ambidexterity* is according to O'Reilly and Tushman (2013) the ability to organize and manage both exploration and exploitation in the same organization, even though conflicting structures, incentives etc. will exist. The reason for introducing the classical innovation theory into the study of co-production is argued to be relevant, since the New Public Management paradigm was based on the premise that public service organizations existed in stable environments (Jensen and Krogstrup, 2017) and therefore would prosper by using a 'mechanistic management system' – exploitation (Burns and Stalker, 1961). However, the shift to adapt to environments that are in flux requires an 'organic management system' (Raisch and Birkenshaw, 2008). In these organic management systems, projects (or even portfolios of projects) seeking exploration is processed by the public service providers. Or they are addressing the ambidextrous management perspective by designing dedicated innovation management and innovation project teams formally in their organizations, so they can respond more agilely to the societal and demographic changes that emerge (see e.g. Brix, 2017). To build the foundation for an integrative framework between co-production and evaluation, we propose Figure 1 below that enables us to differentiate between the different arenas in which co-production can take place. The two-by-two matrix proposed below is created by relying on the degree of 1) citizen participation in, and 2) citizen influence on the co-production project.

Figure 1: Arenas for co-production



Source: Authors' own development; inspired by Krogstrup and Mortensen (2017)

These exemplar arenas presented above are imperative to consider in the context of evaluation, because the conditions, the premises and the mandates given to both the public service provider and to the citizen are different. Therefore, we argue that when selecting the relevant evaluation model, we need to take the arena for co-production into account, because the contextual conditions in these arenas will be different. In the following, we provide an example of a public service provided in each of the arenas presented above. If there is low degree of citizen participation and low degree of citizen influence, then the front-end staffer is doing *casework*, and not co-production. An example could be that a citizen contacts the municipal employee to ask for financial resources to pay for (part of) a visit at the dentist, or that an elderly contacts the municipality to get help with transportation from her home and to a doctor's appointment. An example of having high citizen influence and low degree of citizen participation, *exploitation*, is when the citizen is asked to select which types of pre-defined services s/he would like to receive e.g. make it easier to live with a chronic disease. The citizen can select from a pre-given list in consultation with a medical doctor, but s/he cannot invent new services. An example of high citizen participation and low citizen influence, *exploration*, can be an innovation project, where the citizens are invited to workshops with the purpose of generating new ideas for how to rethink public services, but where the front-end staffer or other municipal employees have decision power to select which ideas that are to be further developed (see. e.g. Brix & Peters, 2015; Brix, 2017). The fourth context for public service development and innovation is labelled *ambidextrous*. Here there is both high citizen participation and the citizen has high degree of influence. An example of co-production in this context can be found in the municipality of Ikast-Brande in Denmark. Here a group of citizens created an initiative with support from the municipal employees that had the purpose of increasing the degree of collaboration between citizens, private companies and the different leisure associations. The citizens who took the initiative collaborated on equal basis with the employee staff and created an 'volunteer stock-exchange' (in Danish: FrivilligBørsen) in which everyone interested in making e.g. barter agreements or something alike could meet and network. (CFSA, 2013). By having exemplified the three types of arenas for co-production, we proceed to discuss how it is possible to evaluate the outcomes of the types of co-production that occur in the ambidextrous arena. The other arenas – exploration and exploitation – are rather straightforward to evaluate, and therefore we wish to address the more complex and uncertain arena of ambidexterity. In the following we continue to discuss the premises of the ambidextrous approach to co-production to identify the framing conditions for evaluation.

Linking evaluation research to co-production in the ambidextrous arena

The ambidextrous arena for co-production can be recognized by three characteristics which are important to consider when deciding upon evaluation methodologies. The first characteristic is, that there are many stakeholders involved in the co-production process, which means that the contents of the process constantly emerge and change as the interaction between the stakeholders occurs (Peronard and Brix, 2017). The second characteristic pertaining to this arena, is that co-production is not limited to one specific context but instead the co-production processes take place in many different contextual sub-arenas in which the stakeholders interact. This leads us to the third characteristic of co-production in the ambidextrous arena: there is no causal relationship between the intervention and the outcome, because multiple variables can effect the relationship. Hence, we argue that co-production in the ambidextrous arena is a complex, social problem and that it cannot be limited to a technical one. Since we need to utilize an evaluation model – or different evaluation models – that take(s) into account the three characteristics presented above, and because our goal is to study the outcome of the complex phenomena ‘co-production’, it is imperative that we identify (a) suitable evaluation model(s) that enable(s) us in reaching our goal. In the following we therefore propose and discuss how the outcomes of prospective co-production projects can be evaluated.

Attribution versus contribution evaluation

According to effect evaluation the classical dichotomy is made between attribution analysis and contribution analysis. Attribution analysis is used to assess to which extent a given intervention has attributed to the outcome. *Has the intervention caused the outcome? To what extent has the intervention caused the outcome?* The randomized controlled trial (RCT) method is golden standard in attribution analysis (Bichman and Reich 2009). RCT regards problems as simple and not as complex, and as technical instead of social. An attribution analysis is therefore unsuitable to apply when evaluating complex programs consisting of many interacting variables, collaborative interventions and unpredictability (Mayne 201; Forss and Schwartz, 2011; Patton, 2112). Hence, we stress that the attribution analysis which has dominated the new evidence wave does not apply to the criteria for evaluating outcomes of co-production in the ambidextrous arena we presented above. Instead, we suggest that evaluators of the outcomes of co-production take point of departure from a contribution analysis perspective. Contribution analysis is defined as an approach addressing cause-effect questions build on the idea of using an intervention theory of change to infer causation (Leeuw, 2012). Key questions in the contribution evaluations are: *Has the intervention made any*

difference for the subject(s)?, Has the intervention influenced the observed result?, How big a difference has the intervention had?, How big is the contribution? A contribution analysis therefore produces a contribution story that represents the evidence and other influences on program outcomes: a major part of that story may tell about behavioral changes that intended beneficiaries have made as a result of the intervention (Patton, 2012). A result of the contribution analysis does not represent the ultimate truth. It has to be regarded as a sufficient conclusion concerning the extend – and the reason(s) why – the intervention has contributed with a given outcome. The problem is that it is difficult to define an expected outcome in the ambidextrous arena, because the desired outcome will first be identified during the co-production process. Below we review evaluation models pertaining to contribution analysis.

Evaluation models pertaining to contribution analysis

In the beginning of the 1970's responsive evaluation was developed and promoted by Robert Stake who got inspired by the Nordic welfare model during a research semester in Uppsala in Sweden during his doctoral studies. This humanistic orientation to public welfare inspired him to develop an alternative to RCT; being responsive evaluation. Responsive evaluation is known for its focus on case studies. Here the evaluator interacts with the field, and both qualitative and quantitative methods can be used to determine outcomes. Robert Stake's work has inspired many evaluations theorists. His approach to evaluation has been developed since the 1970's and it has been expanded under different headlines such as 'interactive evaluation' (Rodrigues, 2012; King and Stevahn, 2013; Owen and Rogers, 1999), 'collaborative evaluation' (King and Stevahn, 2013), 'utilization-based evaluation' (Patton 2012), and 'dialogue evaluation' (Vedung, 2010). It is though a general misconception that the responsive evaluation approach is a process evaluation approach. Evaluation based on this approach can also be used to evaluate the outcomes of interventions. Inspired by Stake's work, Guba and Lincoln (1989) developed what they coined the 'fourth generation evaluation'. Fourth generation evaluation is recognized by its long pre-evaluation focus in which the evaluation questions are negotiated among the evaluation's stakeholders. Another important representative inspired by responsive evaluation is Michael Quinn Patton. He focused by his 'utilization-based evaluation' on the use of evaluation which was increased by involving stakeholders in the evaluation to get ownership to the evaluation (Patton, 2012). The ownership was created by collaboration or participation during all steps in the evaluation process from setting up the evaluation question to the final report. Patton focused on the changes that occurred during the

evaluation process created by all the users of the evaluation. The formative, emerging elements, represent important aspects of Patton's evaluation approach (ibid). Another important representative of this approach to evaluation, is 'empowerment evaluation', which focuses on the evaluation process as a source of empowering the users by participating in the evaluation process Fetterman, 2001; 2013; Fetterman, Kaftarian and Wandersman, 2015). The latest we have heard from this approach to evaluation is Shulha et al. (2016), focusing on 'collaborative evaluation'. Shula et al. (2016) use a multimethod perspective to identify and construct guiding principles for collaborative evaluation and in this work they have involved 320 responses primary from members of American Evaluation Association. In their study they identify eight guiding principles that are important when planning, conducting and ending a collaborative evaluation approach (see also Shulha 2015; Patton 2012, King and Stevahn 2013).

These principles are:

- Clarify motivation for collaboration
- Foster meaningful relationships
- Develop a shared understanding of the program
- Promote appropriate participatory processes
- Monitor and respond to resource availability
- Monitor evaluation progress and quality
- Promote evaluative thinking
- Follow through to realize use – which changes is important?

We find it important to further investigate these principles stemming from collaborative evaluation in the context of co-production to investigate if they can enable us in responding to the OECD's (2011) request for increased focus on co-evaluation and not least the focus on the outcomes of co-production projects.

Collaborative evaluation as a method to evaluate to outcomes of co-production?

According to our ambition to present an outcome-oriented evaluation approach, we need to define the evaluation's questions: *Has the co-production intervention had any differences for the subject(s)?, Has the co-production influenced the observed outcome?, How big a difference has the co-production had?* Another question which is important when we want to evaluate the connection between co-production and its outcome is to ask *which mechanisms have caused the outcome?* By pre-defining the evaluation questions instead of letting the stakeholders participate in this work we have already left the concept of collaborative evaluation. That means that we cannot be sure that

there is “a high level of agreement among the stakeholders about program information need” and that “program goals consistently echo across stakeholders” which Shulha et. al (2016) describe as important indicators for a collaborative evaluation. This is however the consequence of mixing contribution evaluation with a collaborative evaluation approach. However, in the spirit of co-production this evaluation does not operate with pre-defined criteria for outcome of the co-production process; the outcome criteria are negotiated as the first stage of the evaluation according to the guiding principles of collaborative evaluation. This can be seen as a way to handle the contradictory conditions that exist, when we want to measure outcome without having any outcome criteria. It is obvious that the stakeholders involved in the co-production process participate in an inclusive evaluation process. However, we wonder if the eight guidelines and principles proposed by Shulha et al. (2016) also could act as guidelines for co-production in a way where the collaborative evaluation process both is able to evaluate the outcome of the co-production process, to identify the generative mechanisms, and also to build capacity to co-produce on an individual and organizational level? Capacity Building in this regard means that individuals just as organizations increase their ability to co-produce by increasing the skills to collaborate, identifying needs for change and development and getting new competences to fulfil individual as organizational goals (Krogstrup, 2017). The triggers for learning is partly “learning by doing” in the evaluation process (Cousins et al. 2014) or processes inspired by collaborative evaluation or by building knowledge about generative mechanism. Knowing the generative mechanisms, the idea is not to copy co-production processes e.g. from Denmark to Sweden and by this expect the same outcome – this idea of direct transfer lingers in the attribution evaluation paradigm, and not the one of co-production in complex, human problems (Kildedal and Krogstrup, 2010; Krogstrup, 2016). An idea could be to rely on theory-based evaluation and utilize the eight principles of collaborative evaluation to support the development of the change theory. This is argued, because if we are able to reproduce the generative mechanisms and in collaborative processes generate knowledge about reasonable outcome criteria, we can build a strong local change-theory to develop the co-production concept. Further this will in a broader sense probably result in the possibility building up and accumulate knowledge concerning the conditions under which co-production will occur: what has an effect and what does not. The point is that co-production activities will vary across different context but the measure for quality of the process is whether the generative mechanisms are active: “Principle for (co-production) are built from lessons that are based on evidence about how to accomplish some desired result” (Patton 2015).

DISCUSSION

Theorists traditionally recommend that existing data has to be used to conduct a contribution analysis (Leeuw 2012; Toulemonde 2011). However, our review above determines that only scarce empirical evidence exists concerning the effects of co-production in the context of public service organizations (see. e.g. Voorberg et al. 2012). Interestingly, our review indicates that co-production as a new organizational recipe is widely accepted as being the new panacea without having the evidence for backing up the claims for the effects. We claim that the arena for co-production has to be used to decide on the choice of evaluation model. The arena for co-production sets the boundary for the decision mandate, the premises and framing conditions of the evaluation. And these will be different from municipality to municipality, country to country (Dahler-Larsen, 2013). In this study we delve into the discussion of the ambidextrous arena, which is a complex context constituted by three criteria: 1) there are many stakeholders involved, 2) there are multiple sub-arenas for interaction, and 3) there is no causal relationship between the intervention and the outcome. This enables us to determine, that contribution analysis-oriented evaluation models can be appropriate to study outcomes of co-production processes. The big challenge in this kind of evaluation is to describe the change theory (Krogstrup, 2016). The change theory emerges and converts during the co-production process involving multiple stakeholders in many contextual arenas. A change theory has to be well-argued which means that it has to be plausible and substantiated with theory and/or empirically founded facts (Leeuw, 2012; Funell and Rogers, 2011). A way to handle this complex co-production process in the ambidextrous arena is to break down the change theory in manageable and definable sub-theories, so these can be pursued as partial pieces of the synthesized change theory (Pawson, 2006). The validation of the identified outcomes is made as a communicative process between the participating stakeholders. The goal is to reach an agreement about the expected connection between co-production and its outcome. This means that the stakeholders have to agree on the change theory. To cope with this challenge, we propose that the eight principles of collaborative evaluation (Shulha et al. 2016) might be used as guidelines for co-production projects. However, we stress that the users do not have to co-decide whether or not the evaluation has to center on the outcomes. Consequently, we suggest contribution analysis as a qualified evaluation approach, we need to consider the generative mechanisms that can be identified to trigger the outcome(s) of co-production processes (Krogstrup 2016). In this regard it is imperative that scholars and practitioners are clear in the distinction between the description of the mechanisms, the mediators and the moderators in their change theory and that there is a clear logic between their

interlinkages. According to Patton this clear, logical thinking is imperative when creating change theory (2012) “*rigorous thinking was more important than relying on methodological rigor*” (Ibid: 374). Mechanism is defined as a process which connects the cause and the effect and in that sense explains why an outcome occurs when one or more causes exist. Mediators are the intervening variables which potentially influence the co-production project, and vice versa. Moderators refers to the conditions which regulate the causality between co-production and the outcomes/effect of the process (Bamberger et al. 2012; Dahler-Larsen, 2013). When the scholar and/or practitioner has created the change theory, the co-production process can be initiated, continuously evaluated and adjusted to fit the new knowledge emerging from the process, so the generative mechanisms can be identified that are expected to cause the desired outcomes of the process.

IMPLICATIONS

Currently, we and other members of the 'Research Group for Evaluation & Capacity Building' at Aalborg University in Denmark are conducting two large studies concerning co-production with low-socioeconomic, socially marginalized and/or disabled citizens as specific groups of people that are more challenging to co-produce with compared to citizens without special needs.

One of the studies is an Industrial PhD-project funded by the Innovation Fund Denmark. The project studies co-production processes on an individual level (micro level), in the empirical context of service units to disabled citizens. The project has a multiple-case design and the purpose is two-folded. First, the study seeks to explore and evaluate the actual co-production processes, and second, the study seeks to identify the outcomes of these processes. The process evaluation determines the barriers and drivers experienced by the citizens and employees during the implementation process of co-production. The process evaluation is structured by the bottom-up oriented evaluation model BIKVA, where data are produced in focus group interviews (Krogstrup 2013). The aim is to assess whether co-production processes actually occurs and establish different influencing factors like staff and management involvement, engagement, trust, shared meaning, existing organizational values, practices and culture, risk-taking at all organizational levels, employees' reluctance to lose status and control and the need for staff training (OECD 2011: 88, Bovaird & Loeffler 2012, 11-12, Labitzke et al. 2014, Boyle & Harris 2009, Agger & Tortzen 215, Jensen & Krogstrup 2017). A theory based effect evaluation with theories of change will be established with the purpose of generating evidence of the effects of the co-production processes.

The industrial PhD project strives to determine the different change theories that are implicitly used as foundation for different co-production methodologies. The evaluation will explore whether the implementation of co-production leads to the expected effects and outcomes. It will bring clarification to the question *what works for whom under which circumstances*. (Chen 1990, Pawson & Tilley 1997, Dahler-Larsen 2013, Krogstrup 2006, 2016).

In another study, we are currently reviewing more than 130 documents reporting on the use of the BIKVA² model (Krogstrup, 1997a; Dahler-Larsen and Krogstrup, 2003; Krogstrup, 2016). Here we are treating another special issue concerning co-production: the asymmetric relationship between the users and the member(s) of the organization like e.g. front-staffers. These BIKVA-reports present empirical evidence for how public service professionals – in projects ranging from evaluating or developing local housing offers to the use of BIKVA as methodological foundation in work packages in large European projects – have overcome the barriers of involving citizens with different user profiles in the projects (see also Brix, Nielsen and Krogstrup, 2017).

CONCLUSION

Our study echoes the necessity for increasing the amount of empirical studies evaluating the outcomes of co-production initiatives and the need for understanding how to operationalize co-evaluation (Löffler, 2009; OECD, 2011; Voorberg, et al., 2013; Jensen and Krogstrup, 2017). More precisely, our conceptual discussion argue that contribution analysis represented by theory-based evaluation can be the appropriate method to evaluate the effects of co-production initiatives in public service organizations. Since the literature is scarce concerning empirical studies of evaluation concerning co-production, we have to approach the contribution analysis utilizing empirical data. We propose that the principles of collaborative evaluation (Shulha et al. 2016) can be used to frame this data collection. In line with this, we also argue, that the eight collaborative evaluation principles might act as principles for co-evaluation in the context of OECD's (2011) call for future research. Moreover, we argue that the principles of collaborative evaluation can contribute to co-production Capacity Building. Finally, we call for empirical research to back up the arguments provided in present study.

² BIKVA is the acronym for 'BrugerInvolvering i KVAlitetsevaluering' which in English is translated to 'User Participation in Quality Assessment' (See Krogstrup, 1997b)

REFERENCES

- Bason, C. (2010). *Leading public sector innovation. Co-creation for better society*. The Policy Press.
- Bickmann, L. and Reich, S. (2009): Randomized Controlled Trials: A gold Standard with Feet of Clay. I Donaldson, Stewart, Christian A Christie & Melvin Mark (red). What Counts as Credible Evidence in Applied research and Evaluation practice? Thousand Oaks, Sage
- Bovaird, T., & Loeffler, E. (2012). From Engagement to Co-production: The Contribution of Users and Communities to Outcomes and Public Value. *Voluntas*, 23(4), 1119–1138.
<https://doi.org/10.1007/s11266-012-9309-6>
- Boyle, D., & Harris, M. (2009). The challenge of Co-Production - How equal partnerships between professionals and the public are crucial to improving public services. *Public Administration*, 28. Retrieved from http://b.3cdn.net/nefoundation/312ac8ce93a00d5973_3im6i6t0e.pdf
- Brix, J. (2017) Exploring knowledge creation processes as source of organizational learning – A longitudinal study of a public innovation project, *Scandinavian Journal of Management*, 33(2): 113-127
- Brix, J., & Peters, L. S. (2015). Exploring an Innovation Project as a Source of Change in Organization Design. *Journal of Organization Design*, 4(1), 29-43.
- Brix, J., Nielsen, L. K., & Krogstrup, H. K. (2017). The Diffusion and Adoption of BIKVA: A 20-year Impact Case Study, *Research Group for Evaluation & Capacity Building*, Aalborg University
- Brudney, J. L., & England, R. E. (1983). Toward a Definition of the Coproduction Concept. *Source: Public Administration Review*, 43(1), 59–65. Retrieved from <http://www.jstor.org/stable/975300>
- Burns, T.E. and Stalker, G.M. (1961) The Management of Innovation, *University of Illinois at Urbana-Champaign's Academy for Entrepreneurial Leadership Historical Research Reference in Entrepreneurship*. Available at SSRN: <https://ssrn.com/abstract=1496187>
- CFSA (2013) *FrivilligBørs Ikast-Brande – Evalueringsrapport*, Center for Frivillig Socialt Arbejde, Odense, Juli.
- Cousins, B.J., Swee C.C., Elliott, C.J. and Bourgeois I. (2014): *Framing the Capacity to Do and Use Evaluation* In B.J. Cousins & I. Bourgeoise (Eds.) *Organizational capacity to do and use Evaluation*. New Directions for Evaluation, 141 7-23
- Dahler-Larsen, P. (2013): *Evaluering af projekter – og andre ting, som ikke er ting*, Syddansk Universitetsforlag.
- Dahler-Larsen, P. og Krogstrup, H.K. (2003) *Nye veje i evaluering*. Systime Academic
- Fettermann, D.M., Kaftarian, S. and Wandersman, A. (2015): *Empowerment Evaluation*. 2 editions. Sage
- Fettermann, D.M. (2013): *Empowerment Evaluation: Learning to Think like an Evaluator*. In: Alkin, Marvin (red. 2013): *Evaluation Roots - A wider Perspective of Theorists' Views and Influences*. Sage

- Fetterman, D.M. (2001), *Foundation of Empowerment Evaluation*, Sage
- Forss, K. and Schwartz, R. (2011): Introduction. In: Forss, Kim. Mita Marra and Robert Schwartz (2011): *Evaluating the Complex – Attribution, Contribution and Beyond Comparative Policy Evaluation* Volume 18
- Funnell, S.C & P.J. Rogers (2011): *Purposeful Program Theory - Effective Use of Theories of Change and Logic Models*, Jossey-Bass.
- Gouillart, F., & Hallett, T. (2015). Co-creation in Government. *Stanford Social Innovation Review*, 40–48. <https://doi.org/10.1177/1532673X08330635>
- Guba, E.G. and Lincoln, Y. (1989), *Fourth Generation Evaluation*, Sage.
- Hood, C., & Dixon, R. (2013). A model of cost-cutting in government? the great management revolution in uk central government RECONSIDERED. *Public Administration*, 91(1), 114–134. <https://doi.org/10.1111/j.1467-9299.2012.02072.x>
- Jakobsen, M., & Andersen, S. A. (2013). Co-Production and Innovation in Public Service. *Public Administration Review*, 73(5), 704–712. <https://doi.org/10.1111/puar.12094.ver>
- Kildedal, K. og Krogstrup, H.K (2015): *Evidensbaserede koncepters betydning for arbejdet med børn og unge*; Hans Reitzels Forlag
- King, J.A. (2013): *Getting People Involved: The Origen of Interactive Evaluation Practice*. I . I: Alkin, Marvin (red. 2013): *Evaluation Roots - A Wider Perspective of Theorists` Views and Influences*. 2. Edition Sage pp 334-343
- King, A, Jean and Laurie Stevahn (2013): *Interactive Evaluation Practice – Mastering the Interpersonal Dynamics and Program Evaluation*. Sage
- Krogstrup, H. K. (1997a). *Brugerinddragelse og organisatorisk læring i den sociale sektor*. Systime Academic.
- Krogstrup, H. K. (1997b). User participation in quality assessment: a dialogue and learning oriented evaluation method. *Evaluation*, 3(2), 205-224.
- Krogstrup, H. K. (2016). *Evalueringsmodeller*. Hans Reitzel.
- Krogstrup, H.K. og Jensen, J.B. (2017) *Fra New Public Management til New Public Governance*. I Krogstrup, H.K. (eds) (2017) *Samskabelse og Capacity Building i den offentlige sektor*, 1. Udgave, Hans Reitzels Forlag, København
- Krogstrup, H.K. og Mortensen, N.M. (2017) *Opbygning af kapacitet til samskabelse og brugerinvolvering*. I Krogstrup, H.K. (eds) (2017) *Samskabelse og Capacity Building i den offentlige sektor*, 1. Udgave, Hans Reitzels Forlag, København
- Leeuw, F.L. (2012): Linking theory-based evaluation and contribution analyses: Three problems and a few solutions. *Evaluation* 18(3) 348-363

- Levy, R. (2010). New public management: End of an era? *Public Administration Review*, 25, 234-240.
- Löffler, E., (2009). A Future Research Agenda For Co-production Overview paper. LARCI Commissioned Summary Report.
- MandagMorgen (2013) *Radikal velfærdsinnovation – hvordan får vi mere effekt af innovationen?*, Tænk tanken for Radikal Velfærdsinnovation, København
- March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization Science*, 2(1), 71-87.
- Mayne, J. (2011): Contribution Analysis: Addressing Cause and Effect In: Forss, Kim. Mita Marra and Robert Schwartz (2011): Evaluating the Complex – Attribution, Contribution and Beyond Comparative Policy Evaluation Volume 18
- Needham, C., & Carr, S. (2009). Co-production: an emerging evidence base for adult social care transformation. *Research Briefing*, (March), 1–24.
- OECD (2005). *Oslo manual—Guidelines for collecting and interpreting innovation data*. 3rd edition. Organisation for Economic Cooperation and Development, OECD.
- OECD. (2011). *Together for Better Public Services: Partnering with Citizens and Civil Society*. OECD Public Governance Reviews. <https://doi.org/10.1787/9789264118843-en>
- O'Reilly, C. A., & Tushman, M. L. (2013). Organizational ambidexterity: Past, present, and future. *The Academy of Management Perspectives*, 27(4), 324-338.
- Osborne, S. P., Radnor, Z., & Strokosch, K. (2016). Co-Production and the Co-Creation of Value in Public Services: A suitable case for treatment? *Public Management Review*, 18(5), 639–653. <https://doi.org/10.1080/14719037.2015.1111927>
- Ostrom, E. (1996). Crossing the Great Divide : Synergy , and Development. *World Development*, 24(6), 1073–1087. [https://doi.org/10.1016/0305-750X\(96\)00023-X](https://doi.org/10.1016/0305-750X(96)00023-X)
- Ostrom, E., Parks, R. B., Whitaker, G. P. & Percy, S. L. (1978). The public service production process: A framework for analyzing police services. *Policy Studies Journal*, 7, 381–9.
- Owen, J.M & Rogers (1999): Program Evaluation. Form and Approaches. Sage
- Patton, M.Q. (2012a): Essentials of Utilization-focused evaluation; Sage Publication; Minnesota
- Parks, R. B., Baker, P. C., Kiser, L., Oakerson, R., Ostrom, E., Ostrom, V., Percy, S. L., Vandivort, M. B., Whitaker, G. P. & Wison, R. (1981). Consumers as co-producers of public services: Some economic and institutional considerations. *Policy Studies Journal*, 9(7), 1001–11.
- Peronard, J. P., & Brix, J. (in press). Sensemaking in collaborative networks: Creating and balancing activities in open business models. *International Journal of Entrepreneurship and Innovation Management*, XX(y): 000

- Pestoff, V. (2015). Co-production as social innovation in public service. In *Towards a New Local Welfare: Best Practices and Networks of Social Inclusion*.
- Raisch, S., & Birkinshaw, J. (2008). Organizational ambidexterity: Antecedents, outcomes, and moderators. *Journal of Management*, 34(3), 375-409.
- Rodrigues – Campos L (2012):Stakeholder involvement in evaluation: Three Decades of the American Journal of Evaluation, Journal of Multidisciplinary Evaluation Vol 8. Nr 17 pp 57-70
- Stake, Robert E. (1995): The art of case study research. Sage publications
- Stake, R.E. (1990): Responsive Evaluation. I: Walberg J. Herbert and Geneva D. Haertel (eds.): The International Encyclopaedia of Educational Evaluation, Advances in Education. Pergamon Press, pp 75-77
- Stringer, P. (2013). Theory I Stringer, P. *Capacity building for school improvement: Revisited* (Elektronisk udgave ed.). Rotterdam: Sense Publishers.
- Torfin, J., Sørensen, E. and Røiseland, A. (2016) Transforming the Public Section into an Arena for Co-Creation: Barriers, Drivers, Benefits, and Ways Forward, *Administration & Society*, 1-31,
- Toulemonde, J., Carpenter, D. and Raffier, L. (2011): Coping with the Evaluability Barrier: Poverty Impact of European Support at Country Level. In: Forss, Kim. Mita Marra and Robert Schwartz (2011): Evaluating the Complex – Attribution, Contribution and Beyond Comparative Policy Evaluation Volume 18
- Tuurnas, S. (2016). *The Professional Side of Co-Production*. Acta Universitatis Tamperensis 2163
- Vedung, E. (2010), 'Four Waves of Evaluation', *Evaluation*, 16(3), 263-277.
- Verschuere, B., Brandsen, T., & Pestoff, V. (2012). Co-production: The State of the Art in Research and the Future Agenda. *Voluntas*, 23(4), 1083–1101. <https://doi.org/10.1007/s11266-012-9307-8>
- Voorberg, W., Bekkers, V., & Tummers, L. (2013). Embarking On The Social Innovation Journey : A Systematic Review Regarding The Potential Of Co-Creation With Citizens. Paper for the IRSPM conference, Prague 10-12 April 2013. (320090), 1–44.
- Voorberg, W. H., Bekkers, V. J. J. M., & Tummers, L. G. (2014). A Systematic Review of Co-Creation and Co-Production: Embarking on the social innovation journey. *Public Management Review*, 17(July 2014), 1–25. <https://doi.org/10.1080/14719037.2014.930505>